PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/269,118*

AU 1814

DATE: 11/17/94
TIME: 16:35:04

INPUT SET: S1134.raw

> This Raw Listing contains the General
> Information Section and up to the first 5 pages.

```
 1                              SEQUENCE LISTING
 2
 3   (1)      General Information:
 4
 5        (i) APPLICANT: Inouye, Sumiko
 6                       Hsu, Mei-Yin
 7                       Eagle, Susan
 8                       Inouye, Masayori
 9
10       (ii) TITLE OF INVENTION: Prokaryotic Reverse Transcriptase
11
12      (iii) NUMBER OF SEQUENCES: 42
13
14       (iv) CORRESPONDENCE ADDRESS:
15            (A) ADDRESSEE: Weiser & Associates
16            (B) STREET: 230 South Fifteenth Street, Suite 500
17            (C) CITY: Philadelphia
18            (D) STATE: Pennsylvania
19            (E) COUNTRY: U.S.A.
20            (F) ZIP: 19102
21
22        (v) COMPUTER READABLE FORM:
23            (A) MEDIUM TYPE: Floppy disk
24            (B) COMPUTER: IBM PC compatible
25            (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26            (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
27
28       (vi) CURRENT APPLICATION DATA:
29            (A) APPLICATION NUMBER: US 08/269,118
30            (B) FILING DATE: 10-JUN-1994
31            (C) CLASSIFICATION:
32
33     (viii) ATTORNEY/AGENT INFORMATION:
34            (A) NAME: Weiser, Gerard J.
35            (B) REGISTRATION NUMBER: 19,763
36            (C) REFERENCE/DOCKET NUMBER: 377.5888P
37
38       (ix) TELECOMMUNICATION INFORMATION:
39            (A) TELEPHONE: 215-875-8383
40            (B) TELEFAX: 215-875-8394
41
42
43   (2) INFORMATION FOR SEQ ID NO:1:
44
45        (i) SEQUENCE CHARACTERISTICS:
46            (A) LENGTH: 2176 base pairs
```

```
47              (B)  TYPE: nucleic acid
48              (C)  STRANDEDNESS: double
49              (D)  TOPOLOGY: linear
50
51
52         (ix)  FEATURE:
53              (A)  NAME/KEY: CDS
54              (B)  LOCATION: 640..2094
55
56
57         (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:1:
58
59    TCATCCGCGC GGACACCCCC TCCTACGTGC CCCCCGACGC GGAGAGCGGC GTGGAGACGG        60
60
61    TGTACCGCGT TTCCCTGGAT GGTCACCTGG TGGCGGTGGA GTGGGGCCCG CGCACGGGCT       120
62
63    CGCCGCGTCA CCAGCGGCTC TGGTTCGACT CGGATGCGGA AGCCCCCGGA GCCTACTTCG       180
64
65    CGCGCCTCGA GAAGTTGGCG GCTGACGGCT ACATCGACGC GGCCTCGGCA TTGGTCTAAA       240
66
67    CCCTTCAACC ACGGCTCGGC CGCCACGCGC GGCCGGCAGG ACAGGTGCGA CGAACAGACG       300
68
69    ACGACGTGCG CTTCACGCGC GAGCAGCCGA GAGAGGTCCG GAGTGCATCA GCCTGAGCGC       360
70
71    CTCGAGCGGC GGAGCGGCGT TGCGCCGCTC CGGTTGGAAT GCAGGACACT CTCCGCAAGG       420
72
73    TAGCCTGTTC TTGGCTCTCT CCCTCCTAGG CACTACGGCC AGGGTGGGTA GCGGAGCCAA       480
74
75    CGACGCCACC GCCGTTTACC CACCCCGGCC GTAGTGCCTA GGAGGGGAGA GCCGGTGAGG       540
76
77    CTACCGTGCC CCAGGTAAGA TGGTGGTGCT TTCCCGGCCT CCGTCGACTG CTCGCGCCAT       600
78
79    GTCCCGTCTT CCATCGCCGC GCCCGCCCAA GGTGCAGAC ATG ACC GCC AGG CTG         654
80                                                   Met Thr Ala Arg Leu
81                                                    1                 5
82
83    GAC CCG TTC GTC CCC GCA GCT TCG CCG CAG GCC GTG CCC ACG CCC GAG       702
84    Asp Pro Phe Val Pro Ala Ala Ser Pro Gln Ala Val Pro Thr Pro Glu
85                  10                  15                  20
86
87    CTC ACC GCT CCG TCG TCA GAC GCG GCC GCG AAG CGT GAA GCC CGC CGG       750
88    Leu Thr Ala Pro Ser Ser Asp Ala Ala Ala Lys Arg Glu Ala Arg Arg
89              25                  30                  35
90
91    CTC GCG CAC GAA GCG TTG CTC GTC CGC GCG AAG GCC ATC GAC GAA GCG       798
92    Leu Ala His Glu Ala Leu Leu Val Arg Ala Lys Ala Ile Asp Glu Ala
93              40                  45                  50
94
95    GGC GGC GCC GAC GAC TGG GTG CAG GCG CAG CTC GTC TCC AAG GGG CTC       846
96    Gly Gly Ala Asp Asp Trp Val Gln Ala Gln Leu Val Ser Lys Gly Leu
97          55                  60                  65
98
99    GCG GTC GAG GAC CTG GAC TTC TCC AGC GCC TCC GAG AAG GAC AAG AAG       894
```

INPUT SET: S1134.raw

```
100    Ala Val Glu Asp Leu Asp Phe Ser Ser Ala Ser Glu Lys Asp Lys Lys
101     70              75              80              85
102

103    GCC TGG AAG GAG AAG AAG AAG GCC GAG GCC ACC GAG CGC CGC GCG CTG    942
104    Ala Trp Lys Glu Lys Lys Lys Ala Glu Ala Thr Glu Arg Arg Ala Leu
105                     90              95             100
106

107    AAG CGT CAG GCG CAC GAG GCG TGG AAG GCC ACG CAC GTG GGC CAC CTG    990
108    Lys Arg Gln Ala His Glu Ala Trp Lys Ala Thr His Val Gly His Leu
109                    105             110             115
110

111    GGC GCG GGC GTG CAC TGG GCG GAG GAC CGC CTG GCC GAC GCG TTC GAC    1038
112    Gly Ala Gly Val His Trp Ala Glu Asp Arg Leu Ala Asp Ala Phe Asp
113                    120             125             130
114

115    GTG CCC CAC CGC GAG GAG CGC GCC CGG GCC AAC GGC CTG ACG GAG CTG    1086
116    Val Pro His Arg Glu Glu Arg Ala Arg Ala Asn Gly Leu Thr Glu Leu
117            135             140             145
118

119    GAC TCC GCG GAG GCG CTG GCC AAG GCG CTG GGG CTG AGC GTC TCC AAG    1134
120    Asp Ser Ala Glu Ala Leu Ala Lys Ala Leu Gly Leu Ser Val Ser Lys
121    150             155             160             165
122

123    CTC CGC TGG TTC GCG TTC CAC CGG GAG GTC GAC ACG GCC ACG CAC TAC    1182
124    Leu Arg Trp Phe Ala Phe His Arg Glu Val Asp Thr Ala Thr His Tyr
125                    170             175             180
126

127    GTG AGC TGG ACC ATT CCG AAG CGG GAC GGC AGC AAG CGC ACG ATT ACG    1230
128    Val Ser Trp Thr Ile Pro Lys Arg Asp Gly Ser Lys Arg Thr Ile Thr
129                    185             190             195
130

131    TCC CCC AAG CCT GAG CTG AAG GCA GCG CAG CGC TGG GTG CTG TCC AAC    1278
132    Ser Pro Lys Pro Glu Leu Lys Ala Ala Gln Arg Trp Val Leu Ser Asn
133            200             205             210
134

135    GTC GTG GAG CGG CTG CCG GTC CAC GGC GCC GCC CAC GGC TTC GTG GCG    1326
136    Val Val Glu Arg Leu Pro Val His Gly Ala Ala His Gly Phe Val Ala
137        215             220             225
138

139    GGA CGC TCC ATC CTC ACC AAC GCG CTG GCC CAC CAG GGC GCG GAC GTC    1374
140    Gly Arg Ser Ile Leu Thr Asn Ala Leu Ala His Gln Gly Ala Asp Val
141    230                 235             240             245
142

143    GTG GTC AAG GTG GAC CTC AAG GAC TTC TTC CCC TCC GTC ACC TGG CGC    1422
144    Val Val Lys Val Asp Leu Lys Asp Phe Phe Pro Ser Val Thr Trp Arg
145                    250             255             260
146

147    CGG GTG AAG GGC CTG TTG CGC AAG GGC GGC CTG CGG GAG GGC ACG TCC    1470
148    Arg Val Lys Gly Leu Leu Arg Lys Gly Gly Leu Arg Glu Gly Thr Ser
149                    265             270             275
150

151    ACG CTG CTG TCC CTC CTC TCC ACG GAA GCG CCG CGG GAG GCG GTC CAG    1518
152    Thr Leu Leu Ser Leu Leu Ser Thr Glu Ala Pro Arg Glu Ala Val Gln
```

RAW SEQUENCE LISTING
PATENT APPLICATION   *US/08/269,118*

INPUT SET: S1134.raw

```
153                280                      285                     290
154
155    TTC CGC GGC AAG CTC CTG CAC GTC GCC AAG GGC CCG CGC GCC CTG CCC      1566
156    Phe Arg Gly Lys Leu Leu His Val Ala Lys Gly Pro Arg Ala Leu Pro
157        295                      300                     305
158
159    CAG GGC GCC CCC ACG TCG CCC GGC ATC ACC AAC GCG CTC TGC CTG AAG      1614
160    Gln Gly Ala Pro Thr Ser Pro Gly Ile Thr Asn Ala Leu Cys Leu Lys
161    310                      315                     320                 325
162
163    CTC GAC AAG CGG CTG TCC GCC CTC GCG AAG CGG CTG GGC TTC ACC TAC      1662
164    Leu Asp Lys Arg Leu Ser Ala Leu Ala Lys Arg Leu Gly Phe Thr Tyr
165                      330                     335                 340
166
167    ACG CGC TAC GCG GAC GAC CTG ACC TTC TCC TGG ACG AAG GCG AAG CAG      1710
168    Thr Arg Tyr Ala Asp Asp Leu Thr Phe Ser Trp Thr Lys Ala Lys Gln
169                345                      350                     355
170
171    CCC AAG CCG CGG CGG ACG CAG CGT CCC CCC GTC GCG GTC CTC CTG TCT      1758
172    Pro Lys Pro Arg Arg Thr Gln Arg Pro Pro Val Ala Val Leu Leu Ser
173        360                      365                     370
174
175    CGC GTC CAG GAA GTG GTG GAG GCG GAG GGC TTC CGC GTG CAC CCG GAC      1806
176    Arg Val Gln Glu Val Val Glu Ala Glu Gly Phe Arg Val His Pro Asp
177        375                      380                     385
178
179    AAG ACG CGC GTC GCC CGC AAG GGC ACG CGG CAG CGG GTC ACC GGG CTC      1854
180    Lys Thr Arg Val Ala Arg Lys Gly Thr Arg Gln Arg Val Thr Gly Leu
181    390                      395                     400                 405
182
183    GTC GTG AAT GCG GCG GGC AAG GAC GCG CCC GCG GCC CGA GTC CCG CGC      1902
184    Val Val Asn Ala Ala Gly Lys Asp Ala Pro Ala Ala Arg Val Pro Arg
185                410                      415                     420
186
187    GAC GTC GTC CGC CAG CTC CGC GCC GCC ATC CAC AAC CGG AAG AAG GGC      1950
188    Asp Val Val Arg Gln Leu Arg Ala Ala Ile His Asn Arg Lys Lys Gly
189                425                      430                     435
190
191    AAG CCG GGC CGC GAG GGC GAG TCG CTC GAG CAG CTC AAG GGC ATG GCC      1998
192    Lys Pro Gly Arg Glu Gly Glu Ser Leu Glu Gln Leu Lys Gly Met Ala
193        440                      445                     450
194
195    GCC TTC ATC CAC ATG ACG GAC CCG GCC AAG GGC CGC GCC TTC CTG GCT      2046
196    Ala Phe Ile His Met Thr Asp Pro Ala Lys Gly Arg Ala Phe Leu Ala
197        455                      460                     465
198
199    CAG CTC ACG GAG CTC GAG TCC ACG GCG AGC GCC GCT CCG CAG GCG GAG      2094
200    Gln Leu Thr Glu Leu Glu Ser Thr Ala Ser Ala Ala Pro Gln Ala Glu
201    470                      475                     480                 485
202
203    TGACGCTCAG CGCGCGTCCG TCGCCGACGT GCCGCGCGCC AGCAACGCCG CATTCAGCAA      2154
204
205    CTCCGTCAGC CGGCGCGGGT AC                                               2176
```

INPUT SET: S1134.raw

```
206
207
208    (2) INFORMATION FOR SEQ ID NO:2:
209
210        (i) SEQUENCE CHARACTERISTICS:
211            (A) LENGTH: 263 amino acids
212            (B) TYPE: amino acid
213            (D) TOPOLOGY: linear
214
215       (ii) MOLECULE TYPE: protein
216
217
218
219       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
220
221        Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro
222        1               5                   10                  15
223
224        Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met
225                    20                  25                  30
226
227        Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn
228                    35                  40                  45
229
230        Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys
231            50                  55                  60
232
233        Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu
234        65                  70                  75                  80
235
236        Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser
237                    85                  90                  95
238
239        Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp
240                    100                 105                 110
241
242        Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn
243                    115                 120                 125
244
245        Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp
246            130                 135                 140
247
248        Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu
249        145                 150                 155                 160
250
251        Pro Phe Lys Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp
252                    165                 170                 175
253
254        Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys
255                    180                 185                 190
256
257        Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro
258                    195                 200                 205
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/08/269,118*

INPUT SET: S1134.raw

Line	Error	Original Text